EFFICIENT RISK MANAGEMENT WITH MONTE CARLO

Luca Capriotti
GMAG Global Modelling and Analytics Group
Credit Suisse Investment Banking Division

CAP Workshop on Derivative Securities and Risk Management
Columbia University - November 9th, 2007
Outline:

I - Speeding up Monte Carlo:

- Monte Carlo and Statistical Uncertainties
- Least Squares Importance Sampling (LSIS) and stratification (LSIS+)
  - Examples: European/Asian Options
  - Application to the Libor Market Model

II - Speeding up Risk with Monte Carlo:

- Likelihood Ratio Method
  - Copula-based models
  - Variance Reduction techniques
- Adjoint methods
Monte Carlo Sampling

\[
V = E_P[G(Z)] = \int_D dZ \ G(Z) \ P(Z)
\]

“Crude” MC:

\[
V \approx \frac{1}{N_p} \sum_{i=1}^{N_p} G(Z_i) \pm \frac{\sum}{\sqrt{N_p}} \quad Z_i \sim P(Z)
\]

Statistical Uncertainty

Variance

\[
\Sigma^2 = E_P \left[ G(x)^2 \right] - E_P \left[ G(x) \right]^2
\]

\[
\Sigma^2 \approx \frac{1}{N_p} \sum_{i=1}^{N_p} \left( G(Z_i) - \bar{V} \right)^2
\]

\[
N_p (\text{Given Stat. Error}) \propto \Sigma^2
\]
Importance Sampling

\[
\int_D dZ ~ G(Z) ~ P(Z) = \int_D dZ ~ \frac{G(Z)P(Z)}{\tilde{P}(Z)} \tilde{P}(Z)
\]

Simple Identity

\[
V \sim \tilde{V} = \frac{1}{N_p} \sum_{i=1}^{N_p} W(Z_i) G(Z_i) \quad Z_i \sim \tilde{P}(Z)
\]

Sampling Distribution

\[
W(Z) = \frac{P(Z)}{\tilde{P}(Z)}
\]

Likelihood Ratio

\[
\tilde{\Sigma}^2 = \int_D dZ ~ (W(Z) G(Z) - V)^2 \tilde{P}(Z)
\]

Variance

IS: Choose the new probability density in order to decrease Variance

---

Luca Capriotti – Efficient Risk Management with Monte Carlo
Zero Variance Property

\[ P_{opt}(Z) = \frac{1}{V} G(Z) P(Z) \]

\[ W(Z) = \frac{P(Z)}{\tilde{P}(Z)} = \frac{V}{G(Z)} \]

\[ \tilde{V} \approx \frac{1}{N_p} \sum_{i=1}^{N_p} W(Z_i) G(Z_i) = \frac{1}{N_p} \sum_{i=1}^{N_p} V \]

Too bad I don’t know \( V \) …..

But I can still try to find a Sampling Distribution that is as close as possible to the Optimal one.
Trials Sampling Densities

\[ \tilde{P}_\theta(Z) \]

Set of Optimization Parameters

Optimization Problem:

\[ \tilde{\Sigma}^2 = \int dZ \ (W(Z)G(Z) - V)^2 \tilde{P}(Z) \]

Original Measure

\[ \tilde{\Sigma}_\theta^2 = E_P \left[ W_\theta(Z)G^2(Z) \right] - E_P \left[ G(Z) \right]^2 \]
Least Squares Importance Sampling (LSIS)

Minimize the Variance

\[ \bar{\Sigma}_\theta^2 = E_P \left[ W_\theta(Z) G^2(Z) \right] - E_P \left[ G(Z) \right]^2 \]

... or equivalently minimize:

\[ S_2(\theta) = E_P \left[ \left( W_\theta(Z)^{1/2} G(Z) - V_T \right)^2 \right] \]

with Monte Carlo estimator:

\[ \approx \frac{1}{N_p'} \sum_{i=1}^{N_p'} \left( W_\theta(Z_i)^{1/2} G(Z_i) - V_T \right)^2 \]

\[ \sum_{i=1}^{N_p'} \left( y_i - f_\theta(x_i) \right)^2 \]

\[ x_i \rightarrow Z_i \]

\[ y_i \rightarrow V_T \]

\[ f_\theta(x_i) \rightarrow W_\theta(Z_i)^{1/2} G(Z_i) \]

A Least Squares Problem!
Algorithm:

1) Choose a trial probability distribution and an initial value of the parameters $\theta$

2) Generate a suitable number $N_p'$ of replications of the random variables $Z_i$

3) Set:

$$x_i \rightarrow Z_i \quad y_i \rightarrow V_T$$

$$f_\theta(x_i) \rightarrow W_\theta(Z_i)^{1/2}G(Z_i)$$

4) Feed the pairs $(x_i, y_i)$ into a non linear Least Square Fitter (e.g., Levenberg-Marquardt) to determine the optimal $\theta$. 
Least Squares Importance Sampling

Correlated Sampling makes the approach practical

\[ S_2(\theta) \approx \frac{1}{N_p'} \sum_{i=1}^{N_p'} \left( W_\theta(Z_i)^{1/2} G(Z_i) - V_T \right)^2 \]

A limited number of paths is necessary to determine the optimal \( \theta \)

In fact, the configurations \( Z_i \) are fixed. So, the difference between

\[ S_2(\theta) \quad \text{and} \quad S_2(\theta') \]

is much more accurate than the Monte Carlo estimate of each of them.

Luca Capriotti – Efficient Risk Management with Monte Carlo
European Call

\[ G(Z) = e^{-rT} \left( X_0 \exp \left( r - \frac{\sigma^2}{2} \right) T + \sigma \sqrt{T} Z \right) - K \]

\[ P(Z) = (2\pi)^{-1/2} \exp(-Z^2/2) \]

Trial Density

\[ \tilde{P}_{\mu}(Z) = (2\pi)^{-d/2} e^{-(Z-\mu)^2/2} \]

\[ \tilde{P}_{\mu,\sigma}(Z) = (2\pi\tilde{\sigma}^2)^{-1/2} e^{-(Z-\tilde{\mu})^2/2\tilde{\sigma}^2} \]

\[ N'_p \approx 50 \]

<table>
<thead>
<tr>
<th>( \sigma )</th>
<th>( K )</th>
<th>LSIS(( \tilde{\mu} ))</th>
<th>LSIS(( \tilde{\mu}, \tilde{\sigma} ))</th>
<th>RM</th>
<th>GHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>30</td>
<td>104(1)</td>
<td>1700(100)</td>
<td>112(4)</td>
<td>100(1)</td>
</tr>
<tr>
<td>50</td>
<td>7.8(1)</td>
<td>15(1)</td>
<td>7.8(4)</td>
<td>7.8(1)</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>33.5(5)</td>
<td>84(5)</td>
<td>31(2)</td>
<td>33.5(5)</td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>30</td>
<td>16.4(1)</td>
<td>51(1)</td>
<td>16.8(4)</td>
<td>14.8(2)</td>
</tr>
<tr>
<td>50</td>
<td>9.9(5)</td>
<td>27(1)</td>
<td>11(2)</td>
<td>9.9(1)</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>15.6(1)</td>
<td>35(1)</td>
<td>15.2(4)</td>
<td>14.2(1)</td>
<td></td>
</tr>
</tbody>
</table>

Variance Reduction

\[ \text{VR} = \left( \frac{\sigma(\text{Crude MC})}{\sigma(\text{IS})} \right)^2 \]

European Straddle

\[ G(Z) = e^{-rT} X_0 \exp \left( \left( r - \frac{\sigma^2}{2} \right) T + \sigma \sqrt{T} Z \right) - K \]

Straddle

Bimodal Ansatz:

\[ \tilde{P}(Z) = \left( 2\pi \right)^{-d/2} \left[ w_a e^{-\left( Z - \mu_a \right)^2 / 2} + w_b e^{-\left( Z - \mu_b \right)^2 / 2} \right] \quad w_a + w_b = 1 \]
Stratified Sampling

Stratifying a Normal Random Variable

Reducing the Variance by Sampling in a more regular pattern
LSIS + Stratified Sampling (LSIS+)

Too many sample points to fill the space in high dimension $d$!

I can stratify one-dimensional projections!

$$Z^{(i)} = \xi X^{(i)} + (I_d - \xi \xi^t)Y^{(i)}$$

$1$-d Stratified Normal

$\xi \sim \mu(\text{LSIS})$

$N(0, I_d)$

Luca Capriotti – Efficient Risk Management with Monte Carlo
Asian Option with Stratified Sampling:

\[ G(Z) = e^{-rT} \left( \frac{1}{M} \sum_{i=1}^{M} X_i - K \right)^+ \]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>VR (LSIS)</th>
<th>VR (LSIS+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>0.3</td>
<td>45</td>
<td>8.8(1)</td>
<td>950(20)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>9.9(1)</td>
<td>1225(15)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>55</td>
<td>13.6(7)</td>
<td>1900(100)</td>
</tr>
<tr>
<td>64</td>
<td>0.3</td>
<td>45</td>
<td>9.0(2)</td>
<td>1060(30)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>10.3(5)</td>
<td>1290(30)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>55</td>
<td>12.5(5)</td>
<td>1320(100)</td>
</tr>
</tbody>
</table>

Computational Speed-Up of 3 orders of magnitude!

Luca Capriotti – Efficient Risk Management with Monte Carlo
Libor Market Model Setting

Euler Discretization:
\[
\frac{L_i(n+1)}{L_i(n)} = \exp \left[ \left( \mu_i(L(n)) - \frac{1}{2} \| \sigma_i(n) \|^2 \right) h_e + \sigma_i^T(n) Z(n+1) \sqrt{h_e} \right]
\]
\[
\mu_i(L(t)) = \sum_{j=\eta(t)}^i \frac{\sigma_i^T \sigma_j h L_j(t)}{1 + h L_j(t)}
\]
Risk-Neutral Drift

This fits in the general framework:
\[
V = E_P [G(Z)] = \int_D dZ \ G(Z) \ P(Z)
\]
\[
P(Z) = N(0, I_d) \equiv (2\pi)^{-d/2} \ e^{-Z^2/2}
\]
Trial Density
\[
\tilde{P}_{\tilde{\mu}}(Z) = (2\pi)^{-d/2} \ e^{-(Z-\tilde{\mu})^2/2}
\]
Linear parametrization of the drift (knot points)
Caplet

\[ C_h(T_m) = \left( \prod_{i=0}^{m} \frac{1}{1 + hL_i(T_i)} \right) h(L_m(T_m) - K)^+ \]

<table>
<thead>
<tr>
<th>( T_m ) (years)</th>
<th>( K )</th>
<th>( N_k )</th>
<th>LSIS</th>
<th>LSIS+</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>0.04</td>
<td>1</td>
<td>11.4(1)</td>
<td>1349(1)</td>
</tr>
<tr>
<td>1.0</td>
<td>0.055</td>
<td>1</td>
<td>13.3(2)</td>
<td>2300(2)</td>
</tr>
<tr>
<td>1.0</td>
<td>0.07</td>
<td>1</td>
<td>20.2(1)</td>
<td>4126(4)</td>
</tr>
<tr>
<td>2.5</td>
<td>0.04</td>
<td>1</td>
<td>14.0(1)</td>
<td>1189(1)</td>
</tr>
<tr>
<td>2.5</td>
<td>0.055</td>
<td>1</td>
<td>15.5(1)</td>
<td>897(1)</td>
</tr>
<tr>
<td>2.5</td>
<td>0.07</td>
<td>1</td>
<td>18.1(1)</td>
<td>1831(1)</td>
</tr>
<tr>
<td>5.0</td>
<td>0.040</td>
<td>1</td>
<td>12.7(1)</td>
<td>235.2(5)</td>
</tr>
<tr>
<td>5.0</td>
<td>0.060</td>
<td>1</td>
<td>12.5(1)</td>
<td>237.0(5)</td>
</tr>
<tr>
<td>5.0</td>
<td>0.080</td>
<td>1</td>
<td>14.5(1)</td>
<td>193.3(4)</td>
</tr>
<tr>
<td>7.0</td>
<td>0.04</td>
<td>1</td>
<td>7.9(3)</td>
<td>40.0(1)</td>
</tr>
<tr>
<td>7.0</td>
<td>0.055</td>
<td>1</td>
<td>8.5(4)</td>
<td>43.7(1)</td>
</tr>
<tr>
<td>7.0</td>
<td>0.07</td>
<td>1</td>
<td>8.5(4)</td>
<td>40(1)</td>
</tr>
</tbody>
</table>

Speed-ups: 2 - 3 orders of Magnitude

\[ N_p' \approx 100 \]
Swaptions

\[ V(T_n) = \sum_{i=n+1}^{M+1} B(T_n, T_i) h(S_n(T_n) - K)^+ \]

<table>
<thead>
<tr>
<th>( T_n ) (years)</th>
<th>( T_{M+1} )</th>
<th>( K )</th>
<th>( N_k )</th>
<th>LSIS</th>
<th>LSIS+</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>1.5</td>
<td>0.04</td>
<td>3</td>
<td>6.8(3)</td>
<td>35.2(8)</td>
</tr>
<tr>
<td>0.5</td>
<td>1.5</td>
<td>0.055</td>
<td>3</td>
<td>10.5(4)</td>
<td>143(2)</td>
</tr>
<tr>
<td>0.5</td>
<td>1.5</td>
<td>0.07</td>
<td>3</td>
<td>21.2(6)</td>
<td>209(2)</td>
</tr>
<tr>
<td>0.5</td>
<td>2.5</td>
<td>0.04</td>
<td>3</td>
<td>7.0(3)</td>
<td>41.9(9)</td>
</tr>
<tr>
<td>0.5</td>
<td>2.5</td>
<td>0.055</td>
<td>3</td>
<td>9.8(3)</td>
<td>149(2)</td>
</tr>
<tr>
<td>0.5</td>
<td>2.5</td>
<td>0.07</td>
<td>3</td>
<td>18.6(5)</td>
<td>427(2)</td>
</tr>
<tr>
<td>0.5</td>
<td>5.5</td>
<td>0.04</td>
<td>3</td>
<td>6.8(3)</td>
<td>50(1)</td>
</tr>
<tr>
<td>0.5</td>
<td>5.5</td>
<td>0.055</td>
<td>3</td>
<td>8.5(3)</td>
<td>106(1)</td>
</tr>
<tr>
<td>0.5</td>
<td>5.5</td>
<td>0.07</td>
<td>3</td>
<td>12.0(4)</td>
<td>148(1)</td>
</tr>
<tr>
<td>1.0</td>
<td>6.0</td>
<td>0.04</td>
<td>3</td>
<td>8.0(4)</td>
<td>144(2)</td>
</tr>
<tr>
<td>1.0</td>
<td>6.0</td>
<td>0.055</td>
<td>3</td>
<td>8.6(3)</td>
<td>165(2)</td>
</tr>
<tr>
<td>1.0</td>
<td>6.0</td>
<td>0.07</td>
<td>3</td>
<td>12.7(4)</td>
<td>654(3)</td>
</tr>
<tr>
<td>2.0</td>
<td>7.0</td>
<td>0.04</td>
<td>3</td>
<td>9.2(3)</td>
<td>70(1)</td>
</tr>
<tr>
<td>2.0</td>
<td>7.0</td>
<td>0.055</td>
<td>3</td>
<td>9.7(3)</td>
<td>139(1)</td>
</tr>
<tr>
<td>2.0</td>
<td>7.0</td>
<td>0.09</td>
<td>3</td>
<td>13.9(4)</td>
<td>140(1)</td>
</tr>
<tr>
<td>5.0</td>
<td>10.0</td>
<td>0.04</td>
<td>5</td>
<td>7.3(4)</td>
<td>76(1)</td>
</tr>
<tr>
<td>5.0</td>
<td>10.0</td>
<td>0.055</td>
<td>5</td>
<td>7.4(3)</td>
<td>72(2)</td>
</tr>
<tr>
<td>5.0</td>
<td>10.0</td>
<td>0.09</td>
<td>5</td>
<td>7.5(4)</td>
<td>197(2)</td>
</tr>
</tbody>
</table>

Speed-ups: 1 - 2 orders of Magnitude

Luca Capriotti – Efficient Risk Management with Monte Carlo
Straddle

\[ S_{t\mathcal{h}}(T_m) = \left( \prod_{i=0}^{m} \frac{1}{1 + hL_i(T_i)} \right) h |L_m(T_m) - K| \]

<table>
<thead>
<tr>
<th>( T_m ) (years)</th>
<th>( K )</th>
<th>( N_k )</th>
<th>LSIS</th>
<th>LSIS (MM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>0.04</td>
<td>1</td>
<td>2.0(2)</td>
<td>11.6(5)</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>1</td>
<td>1.4(1)</td>
<td>6.4(3)</td>
</tr>
<tr>
<td>1.0</td>
<td>0.055</td>
<td>1</td>
<td>1.1(1)</td>
<td>6.1(3)</td>
</tr>
<tr>
<td>1.0</td>
<td>0.06</td>
<td>1</td>
<td>1.0(1)</td>
<td>4.0(2)</td>
</tr>
<tr>
<td>1.0</td>
<td>0.07</td>
<td>1</td>
<td>1.1(1)</td>
<td>2.6(1)</td>
</tr>
<tr>
<td>5.0</td>
<td>0.04</td>
<td>1</td>
<td>6.8(4)</td>
<td>24.6(8)</td>
</tr>
<tr>
<td>5.0</td>
<td>0.05</td>
<td>1</td>
<td>5.1(3)</td>
<td>21.2(7)</td>
</tr>
<tr>
<td>5.0</td>
<td>0.055</td>
<td>1</td>
<td>4.0(3)</td>
<td>14.8(5)</td>
</tr>
<tr>
<td>5.0</td>
<td>0.06</td>
<td>1</td>
<td>3.5(3)</td>
<td>15.5(6)</td>
</tr>
<tr>
<td>5.0</td>
<td>0.07</td>
<td>1</td>
<td>2.9(2)</td>
<td>15.8(6)</td>
</tr>
<tr>
<td>5.0</td>
<td>0.09</td>
<td>1</td>
<td>2.0(2)</td>
<td>10.0(4)</td>
</tr>
</tbody>
</table>

MM guess provides sizable improvements

Speed-ups: 
1 Order of Magnitude

Bimodal Ansatz

\[ \tilde{P}(Z) = (2\pi)^{-d/2} \left[ w_a e^{-(Z-\mu_a)^2/2} + w_b e^{-(Z-\mu_b)^2/2} \right] \]

\[ w_a + w_b = 1 \]
II - Speeding up Risk with Monte Carlo:

- Likelihood Ratio Method

\[ V(\theta) = \int dx_1 \ldots dx_N G(x_1, \ldots, x_N) P_\theta(x_1, \ldots, x_N) \]

\[ \bar{\theta}_i = \partial_{\theta_i} V(\theta) = \int dx_1 \ldots dx_N G(x) \partial_{\theta_i} P_\theta(x) \times \frac{P_\theta(x)}{P_\theta(x)} \]

\[ \bar{\theta}_i = E[G(X) \Omega_{\theta_i}(X)] \quad \Omega_{\theta_i}(X) = \partial_{\theta_i} \log P_\theta(X) \]

- Calculation is generally Fast
- Does not require regularity condition on the Payoff
- Requires Knowledge of the PDF
- Variance Properties

Luca Capriotti – Efficient Risk Management with Monte Carlo
Gaussian Copula Models

\[ F(x) = \Phi_N \left( \Phi^{-1} (M_1(y_1)), \ldots, \Phi^{-1} (M_N(y_N)) \right); \Sigma \]

\[ P(x) = \phi_N \left( \Phi^{-1} (M_1(x_1)), \ldots, \Phi^{-1} (M_N(x_N)) \right); \Sigma \prod_{i=1}^N \frac{m_i(x_i)}{\phi(\Phi^{-1}(M_i(x_i)))} \]

Market Implied Marginals

\[ m_i(x_i) = \frac{dM(x_i)}{dx_i} \]

\[ \Omega_\theta(x) = \sum_{i=1}^N \partial_\theta \log m_i(x_i) - Z(x)^T (\Sigma^{-1} - I) \partial_\theta Z(x) \]

\[ Z_i = \Phi^{-1}(M_i(x_i)) \]

- Derivatives of the Marginal distributions can be calculated (at worst) by bumping.
\[ \Omega_{\theta}(x) = \sum_{i=1}^{N} \partial_{\theta} \log m_i(x_i) - Z(x)^T (\Sigma^{-1} - I) \partial_{\theta} Z(x) \]

\[ \bar{\theta}_i = E\left[G(X)\Omega_{\theta}(X)\right] \]

- Difficult to say a priori how good or bad will be the Variance of the LRM estimator.
- Forward-related Risks may diverge for small maturities.

Delta Weight 1d LN Model
\[ \Omega = \frac{Z}{X_0 \sigma \sqrt{T}} \quad \Rightarrow \quad E[\Omega] = 0 \]
\[ \lim_{T \to 0} \text{Var}[\Omega] = \infty \]
Variance Reduction Techniques

- Antithetic Variables

\[ \Omega = \frac{Z}{X_0 \sigma \sqrt{T}} \quad \Rightarrow \quad \Omega = \frac{Z - Z}{2X_0 \sigma \sqrt{T}} \equiv 0 \quad \text{Var}[\Omega] = 0 \]

![Chart showing variance reduction techniques](chart.png)

- Exact
- Simple Call

VR(1yr) ~ 7
VR(1wk) ~ 500

\[ K/F = 50\% \]
Variance Reduction Techniques

- Control Variates:
  - Weights: \( E[\Omega] = 0 \)
  - Forwards: \( \partial_\theta E[X_i] = \partial_\theta F_i \)

Vega 6 months 10 Assets Basket Option

Luca Capriotti – Efficient Risk Management with Monte Carlo
II - Speeding up Risk with Monte Carlo:

- Adjoint Techniques

  - First introduced in Computational Finance by Giles & Glasserman, in the 2006 ‘Smoking Adjoints’ paper.

  - The approach can be generalized to Path Dependent options under any multifactor model.

  - The variance of the estimators is essentially the same of the naïve bumped counterparts (after smoothing).

  - Remarkable speed-ups especially when a large number of sensitivities is required.

  - Only one drawback: the implementation does not come for free …
• Path Dependent Multi Asset “best-of” style Option

- Risk with respect to the complete term structure of forwards and vols.
- Portfolio of Bond Options under the Hull-White model

- Risk with respect to the complete term structure of instantaneous fwd rate, the volatility, and the mean reversion speed.
Summary:

I - Speeding up Monte Carlo:

- LSIS - Least-Squares Importance Sampling:
  - Simple Importance Sampling strategy based on a quick LS Optimization.
  - Can be combined with Stratification for further efficiency gains (LSIS+).
  - LSIS can be used with non-Gaussian/multi-modal trial densities.
  - LSIS and LSIS+ can result in computational savings of orders of magnitude.

II - Speeding up Risk with Monte Carlo:

- Likelihood Ratio Method & Variance Reduction techniques:
  - Antithetic Variables solve the divergence of Delta weights.
  - Simple Control Variates can also help to get stable Risk.

- Adjoint methods:
  - Provide a very efficient and general framework for Risk calculation.